Atomic Mass Prediction with Artificial Neural Networks

xuru.org
xuru@xuru.org

September 7, 2008

Abstract

This document describes the development of artificial neural networks for the task of atomic mass prediction, achieving models with a root mean square error as low as 0.631 MeV. An improved semiempirical mass formula with 8 parameters is developed as well.

1 Nuclear structure and the problem of atomic masses

The problem we deal with in this paper is the prediction of the mass of a neutral atom given the number of neutrons N and protons Z it has. Instead of analytic models of the atomic nucleus, we will make use of artificial neural networks (ANN) to give an approximate solution to this problem. However, a basic understanding of the structure of atomic nucleus is necessary for a proper design of these ANNs[1].

At a first glance the problem could seem trivial, for if we know the number of neutrons and protons (and so of electrons) of the atom and the mass of each one of these particles, we could simply multiply and sum to get the total mass. Nevertheless, the equivalence between mass and energy $E = mc^2$ implies that the energy involved in the binding of these constituent particles of the atom is also mass. Since it is necessary to add energy in order to break these bindings and separate completely all the particles that form the atom, the mass of the atom is smaller than the sum of the masses of its constituent particles.

Electromagnetic forces between electrons and the nucleus are weaker than nuclear (and electromagnetic) forces between nucleons[2]: the binding energy of electrons is of some eV or keV per electron, whereas for nucleons it is of some MeV per nucleon.

[1] A better introduction to the structure of the nucleus than our brief explanation in this section can be found in [2].

[2] Nucleons are the constituent particles of the nucleus, both protons and neutrons.

So to a reasonable degree of accuracy, we can neglect the binding energy of electrons and consider that the binding energy of nucleons in the atom is the difference between the mass of its constituent particles separated and the mass of the atom, multiplied by the mass-energy conversion factor c^2,

$$B_{nucleus} = \left(Z\,(m_p + m_e) + N\,m_n - m\right)\,c^2 \tag{1}$$

where m is the mass of the atom, m_p is the mass of the proton, m_e is the mass of the electron, and m_n is the mass of the neutron. Due to the fact that this binding energy of the nucleus[3] increases as atoms of more nucleons are considered, it is more commonly used the binding energy per nucleon,

$$B_{nucleon} = \frac{B_{nucleus}}{Z+N} = \frac{Z\,(m_p + m_e) + N\,m_n - m}{Z+N}\,c^2 \tag{2}$$

So, the task of predicting the mass of the atom can be viewed as the task of predicting the binding energy. Due to its complexity, physicists only can give approximate solutions to this very important problem. These approximate solutions are based mainly on two models of the nucleus: the shell model and the liquid drop model.

The shell model is similar to the shell model of electrons. An approximate nuclear potential is assumed and from it a shell structure of protons and neutrons -which must be considered independently for the Pauli exclusion principle only holds for equal particles- in the nucleus is derived. When a shell is full the nucleus is particularly stable[4], which happens with the so-called "magic numbers" of Z or N equal to 2, 8, 20, 28, 50, 82, 126, 184, 258... When both Z and N are magic numbers the nucleus is exceptionally stable. The shell model explains better than others the properties of light nuclei composed of few nucleons.

In contrast, the liquid drop model explains better the properties of heavy nuclei composed of a lot of nucleons, considering the nucleus as a liquid droplet that can deform, vibrate, rotate... It has special success when explaining nuclear fission, where a heavy, unstable nucleus splits in two.

Considerations of both models -mainly of the second- and simple reasonings about how do vary the volume and surface area of the nucleus with the number of nucleons and its effects in the nuclear and electromagnetic forces -that try to respectively stick together and split the nucleus- lead to the following simple formula for predicting the binding energy, known as the semiempirical mass formula[5], which is commonly used as a model to explain

[3] As we have mentioned, it really also includes the binding energy of electrons, which we should subtract to be very precise.

[4] It is the nuclear equivalent to the chemical stability of noble gases.

[5] There are many variations and refinements on this formula, which shows the interest that exists in improving it. The one shown here is the original proposed by Weizsäcker in 1935, whose predictions have an average error of the order of 1 %.

nuclear properties:

$$B^{smf}_{nucleus} = a_v\,(Z+N) - a_s\,(Z+N)^{2/3} - a_c\,\frac{Z\,(Z-1)}{(Z+N)^{1/3}} - a_{sym}\,\frac{(N-Z)^2}{Z+N} + \delta(Z,N) \tag{3}$$

where

$$\delta(Z,N) = \begin{cases} a_p\,(Z+N)^{-1/2} & \text{for } Z \text{ and } N \text{ even} \\ -a_p\,(Z+N)^{-1/2} & \text{for } Z \text{ and } N \text{ odd} \\ 0 & \text{for } Z+N \text{ odd} \end{cases} \tag{4}$$

and the constants a_v, a_s, a_c, a_{sym} and a_p are obtained by optimization or by multiple linear regression with known atoms, being[6] $a_v = 15.409$ MeV, $a_s = 16.873$ MeV, $a_c = 0.695$ MeV, $a_{sym} = 22.435$ MeV and $a_p = 11.155$ MeV.

Being this the simplest model for the binding energy, any new one should be at least as good as it.

2 The adequacy of ANNs for the task

As mentioned above, physical models only give an approximate solution to the problem. For example, carbon-12 has a binding energy of 92.162 MeV, but the semiempirical mass formula predicts a value of 90.582 MeV. Thus, in this case it has an error of 1.580 MeV. Better models exist, but none is accurate.

However, the task is easy to state: with two inputs and one output -all of them numeric- and having a large set of examples, a model has to be created that allows predicting the output for new input values, and so an ANN can be used to create the model. Physicists would prefer an analytic model, for ANNs act like unintelligible black boxes, but after all if an ANN could make good predictions it would be useful.

Despite of seeming suitable for an ANN, the problem has some particularities that could difficult the task and that must be taken into consideration:

- The fact that physicists have not found a satisfactory solution to the problem suggests that probably it is very difficult to find it, and that an ANN will not do either. However, an ANN could capture heuristics that have passed unnoticed by physicists. This also means that the

[6] Constants shown here have been taken from [1], which is based on the same atomic mass measurements that the used in this work.

ANN will need multiple layers -also suggested by the many nonlinear terms of the semiempirical mass formula- and enough neurons to capture the complex relation that we aim to model, although not in excess because we want an ANN that abstracts instead of memorizing, one with predictable power, which suggests that the amount of information existing in the weights of the network has to be significantly less than the existing in the set of training examples.

- The masses of the atoms that are not very unstable are already measured, and the predictions we are interested in are in the boundary and slightly beyond of this known region. That is, the model will be used to extrapolate. This is risky, for the observed behavior could change in regions different to those known. However, there is not any physical reason to expect significant differences in the behavior of the binding energy near the boundary of the region of measured atoms. Nevertheless, the use of the model to extrapolate could make the errors in the predictions slightly bigger than expected, a problem also shared with the rest of the models like that of the semiempirical mass formula.

- A good choice for the preprocessing of the inputs can be crucial. The values of Z and N have to be normalized to the input range of the neurons, but decomposing these values in others that according to the physical models are meaningful for the problem could improve the result, even if having some redundancy in the inputs. For example, any of these two values Z and N can be decomposed in the number of nucleons in filled shells and the fraction of occupation of the outer incomplete shell. Some postprocessing of the output can be necessary as well, in order to adjust it to the desired range.

3 Atomic mass data

The atomic mass data used in this work is of the evaluation of atomic masses of the year 2003 that can be found in [4, 5].

These experimental data are very accurate, more, by far, than the predictions that any known model can make. However, some values have been discarded because of being estimates and not experimental measurements, so that from the 3179 initial atoms only 2228 have been used in this work.

As our objective is to create a model for the prediction of atomic masses, we have not to care about the binding energy of electrons. The ANN will model both types of binding energies together, that of nucleons and that of electrons[7]. Had we been interested in nuclear masses, we would have to

[7] Subtracting the binding energy of the electrons could facilitate the task to the ANN and get some improvement, but being these energies small, we have not cared about this option.

subtract the small binding energy of the electrons to get as accurate values of the nuclear masses as possible.

4 Preliminary data analysis

As stated above, the measurement errors are very small. The maximum error of the binding energy per nucleon is of 0.053 MeV, and the mean error is of 0.0007 MeV. In comparison, the maximum error of the predictions made with the semiempirical mass formula is of 33.4 MeV, and the mean error is of 2.44 MeV. Therefore, it is not necessary to care about these measurement errors.

Having a visual idea of what we aim to model can help us much. Figure 1 shows the binding energies per nucleon predicted by the semiempirical mass formula, without the last small pairing term δ that as changes of sign with parity only would help to confuse us. The first four terms of the formula correspond to the liquid drop model, and compared with the experimental data shown in figure 2, it can be seen that the prediction is similar; the liquid drop model describes the general behavior of the binding energy. The binding energy per nucleon forms something like an elongated mountain, and in the neighbouring of the line that goes through its top -or in the so-called stability valley- lie stable atoms.

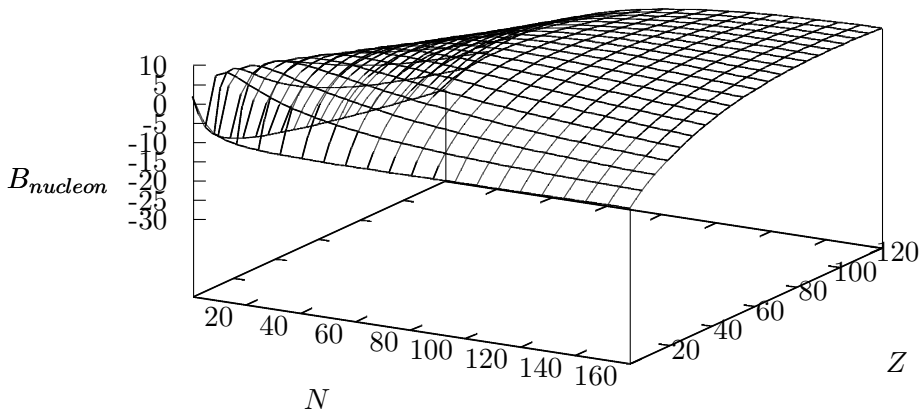

Figure 1: Binding energies per nucleon predicted by the semiempirical mass formula, in MeV.

The form of this surface and the terms of the semiempirical mass formula given in equation (3) suggest that instead of thinking in terms of N and Z it can be better to think in terms of $Z + N$ and $N - Z$.

Figure 3 shows the difference between the prediction of the semiempirical mass formula -including the pairing term δ- and the measured value. The difference does not distribute randomly, but has structure and describes a smooth surface in the N-Z plane, with regions in which it changes little and small transition zones between them where changes are greater. This indicates that still there is a lot that can be modeled and that the semiempirical mass formula can be improved greatly.

These graphics not only show information useful for the modeling task, but also provide direct information on the nuclear structure. Looking at figure 3, we can grasp in a glimpse that the nucleus has in part[8] a shell structure, and we can even predict the number of nucleons in each shell or the "magic numbers" that are given by the transition zones, between which a series of concave surfaces extend.

This can be appreciated better in the projections given in figures 4 and 5, where differences are plotted as a function of Z and N, respectively. Magic numbers are marked by vertical lines, and it can be seen clearly that they coincide with the minima, albeit this is not clear in the case of light atoms,

[8]Most important part of the modeling has been already done by the semiempirical mass formula, i.e. by the liquid drop model, and what we are seeing in figure 3 are the remaining small differences.

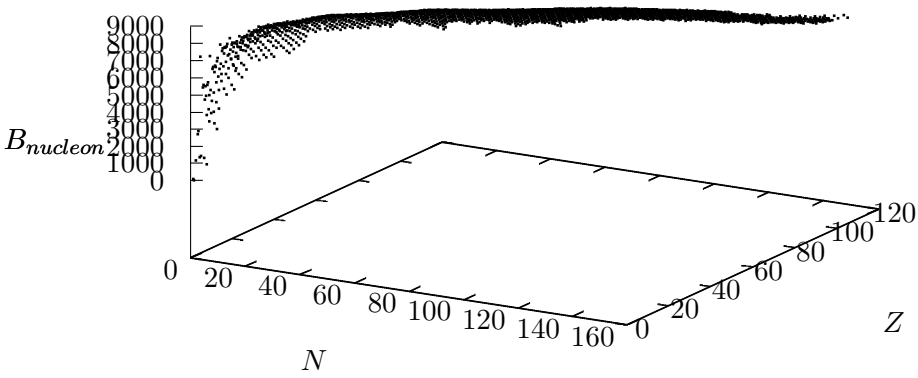

Figure 2: Experimentally measured binding energies per nucleon, in MeV.

where the first three magic numbers cannot be distinguished clearly. This may seem strange, for the shell model is specially successful in explaining the properties of light nuclei, but besides of being few light atoms, some of them -like for example helium-10 with 2 protons and 8 neutrons- have pronounced difference between Z and N, which could disguise the pattern we are seeking.

We should not be surprised by this appearance of the shell model, because the liquid drop model -with the parity[9] included in the pairing term δ- has already been modeled by the semiempirical mass formula.

This means that a way to go beyond the semiempirical mass formula will be to decompose Z and N in such a way to include some information about the shells in the inputs of the ANN. So, inputs should include information about the liquid drop model, about the shell model, and about the parity. This can lead to having some redundancy in inputs, which must be handled carefully.

Finally, figure 6 shows the difference between the prediction of the semiempirical mass formula and the experimental value, as a function of the fraction of occupation of the outer shell of protons and of neutrons, where the general

[9]Parity plays an important role in the shell model. If Z is odd there must be an unpaired proton -respectively an unpaired neutron if N is odd- having a weaker binding. The pairing term δ of the semiempirical mass formula corresponds to considering this effect.

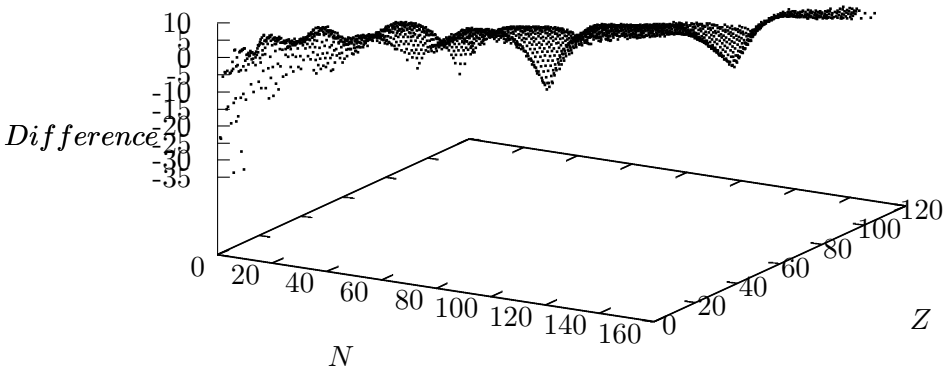

Figure 3: Difference between the binding energy predicted by the semiempirical mass formula and the measured value, in MeV.

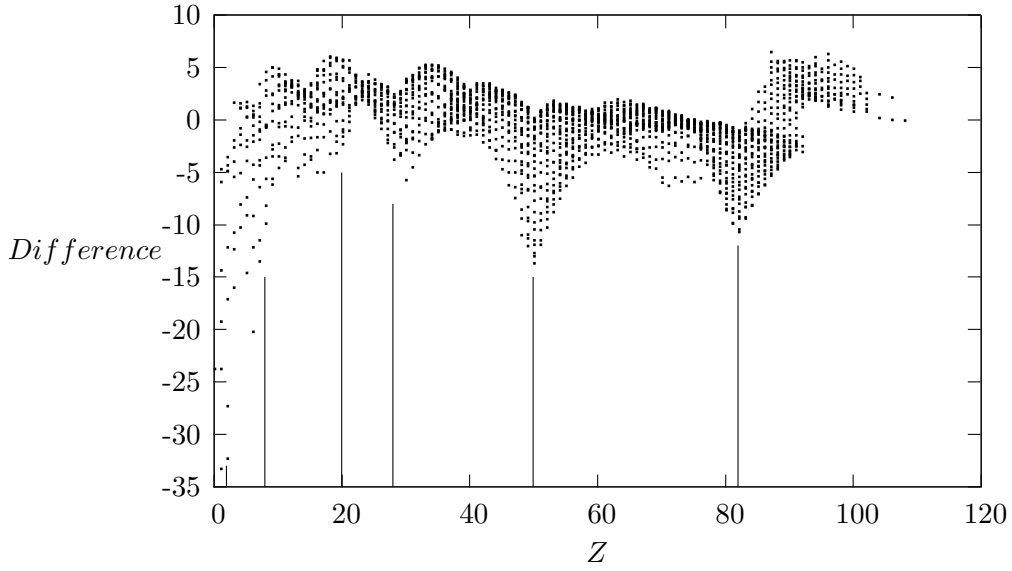

Figure 4: Difference between the binding energy predicted by the semiempirical mass formula and the measured value for different values of Z, in MeV.

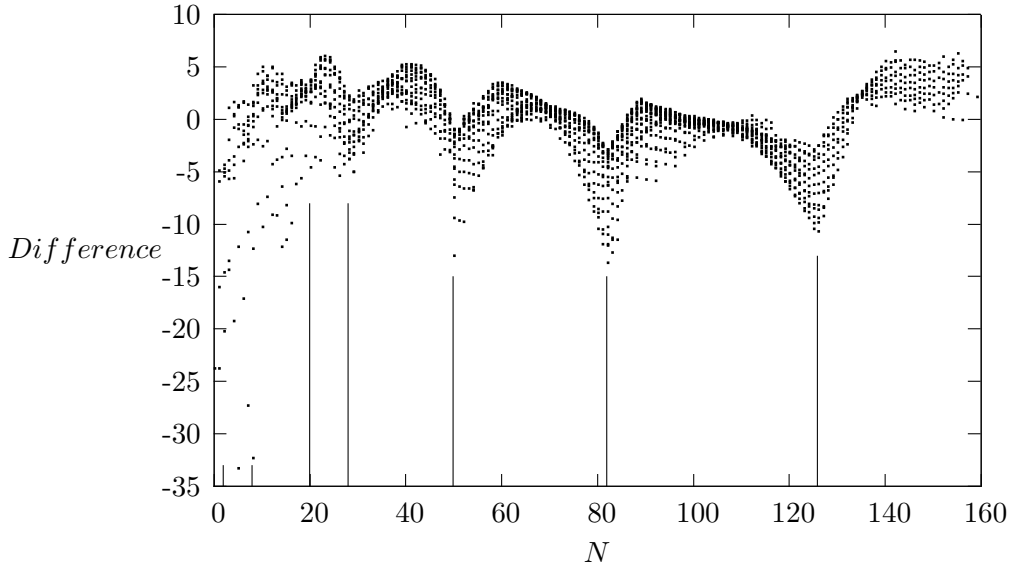

Figure 5: Difference between the binding energy predicted by the semiempirical mass formula and the measured value for different values of N, in MeV.

shape of the concave surfaces mentioned above can be seen.

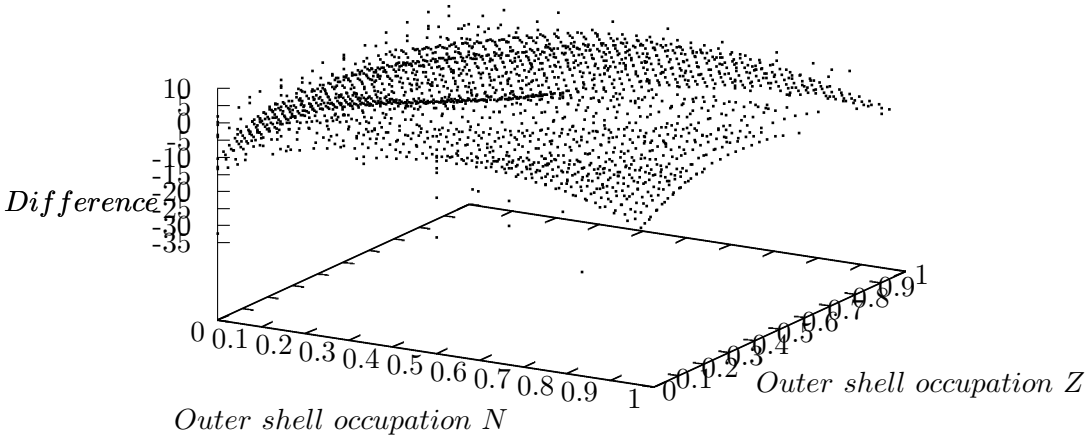

Figure 6: Difference between the binding energy predicted by the semiempirical mass formula and the measured value, as a function of the fraction of occupation of the outer shell of protons and of neutrons, in MeV.

5 Training sets, inputs and outputs

The 2228 examples of the atomic mass data have been randomly divided in three sets: a training set with 60 % of the data (1337 examples), a validation set to prevent overfitting with 30 % of the data (668 examples), and a test set with 10 % of the data (223 examples) for evaluating the performance of the network.

With the results of the preliminary data analysis in mind, the following inputs have been considered:

- The number of protons Z and of neutrons N of the atom.
- The number of nucleons $Z + N$ and the difference between the number of neutrons and of protons $N - Z$.
- The number of filled shells of protons $FS(Z)$ and of neutrons $FS(N)$.
- The number of protons Z_{fs} and of neutrons N_{fs} in filled shells.
- The fraction of occupation of the outer unfilled shell of protons $FO(Z)$ and of neutrons $FO(N)$.

- The parity of Z and of N.
- The value predicted by the semiempirical mass formula or each one of its terms.

Instead of $B_{nucleon}$, the binding energy[10] $B_{nucleus}$ has been considered for the output of the network, for the error in the prediction of the atomic mass is proportional to the error in the prediction of $B_{nucleus}$.

If the range of the input I is between I_{min} and I_{max} while the input of the network I_{net} has to be between a and b, the linear transformation to be used for the preprocessing is the following,

$$I_{net} = a + \frac{I - I_{min}}{I_{max} - I_{min}} (b - a) \tag{5}$$

The ANN will be used to predict the mass of atoms near to those measured. Therefore, we can limit both Z and N to be between 0 and 184, $Z+N$ between 0 and 368, and $N - Z$ between -30 and 100. $FS(Z)$ and $FS(N)$ can be limited to be between 0 and 8, Z_{fs} and N_{fs} between 0 and 184, and $FO(Z)$ and $FO(N)$ will be between 0 and 1. The parity can be regarded as the minimum or the maximum of the neuron input range, depending on whether it is even or odd. Finally, within these limits and disregarding the associated constants, the first term of the semiempirical mass formula will be between 0 and 368, the second between 0 and 52.4, the third between 0 and 4840, the fourth between 0 and 37.4, the fifth between -0.708 and 0.708, and the value of the semiempirical mass formula itself between 0 and 2150 MeV.

On the other hand, if the output O range has to be between 0 and M but the network output O_{net} is between a and b, the necessary postprocessing is

$$O = \frac{O_{net} - a}{b - a} M \tag{6}$$

although in order to calculate the output values that the training examples would have before the postprocessing, the use of the inverse transformation -analogous to (5)- is necessary.

Considering the input and output ranges of the neurons, the values $a = -1$ and $b = 1$ have been used in the above relations.

The error has been measured before the postprocessing instead of after it, but the relation between both is trivial. If the postprocessing is the linear transform $y = ax + b$, O_i is the output of the network for the example i without the postprocessing, V_i is the experimental value that the network

[10] The binding energy really considered has been that of the whole atom, for we have not cared about the small binding energy of electrons.

should give before the postprocessing, $\bar{E}$ is the mean square error before the postprocessing, $\bar{E}^{post}$ is the mean square error after it, E_i is the square error of example i before the postprocessing, and E_i^{post} is the square error of example i after it, then

$$\begin{aligned}\bar{E}^{post} &= \frac{1}{n}\sum_{i=1}^{n} E_i^{post} = \frac{1}{n}\sum_{i=1}^{n}((aO_i+b)-(aV_i+b))^2 = \\ &= \frac{1}{n}\sum_{i=1}^{n}(aO_i-aV_i)^2 = \frac{a^2}{n}\sum_{i=1}^{n}(O_i-V_i)^2 = \\ &= \frac{a^2}{n}\sum_{i=1}^{n} E_i = a^2\bar{E} \end{aligned} \tag{7}$$

Since for Z and N less or equal than 184 the value predicted by the semiempirical mass formula is between 0 and 2150 MeV, we can be confident that the real value of $B_{nucleus}$ will be between 0 and 2200 MeV.

6 Architectures, neuron types and learning algorithm

Having considered the inputs of the previous section, the following ANNs have been developed:

- Option 1: Z and N normalized as inputs, and two feed-forward hidden layers of 9 and 6 neurons, with $2 \cdot 9 + 9 \cdot 6 + 6 \cdot 1 = 78$ connections.
- Option 2: Z and N normalized as inputs, and three feed-forward hidden layers of 7, 8 and 5 neurons, with $2 \cdot 7 + 7 \cdot 8 + 8 \cdot 5 + 5 \cdot 1 = 115$ connections. These first two options will show the modeling possibilities of moderately small ANNs with basic inputs.
- Option 3: Inspired in the information used by the semiempirical mass formula, with $Z + N$, $N - Z$, $Parity(Z)$ and $Parity(N)$ normalized as inputs, and three feed-forward hidden layers of 11, 9 and 5 neurons, with $4 \cdot 11 + 11 \cdot 9 + 9 \cdot 5 + 5 \cdot 1 = 193$ connections. The inclusion of the parities will improve the result. However, since in the ANNs of options 1 and 2 the neurons of the first hidden layer can already have at the input a weighted sum equivalent to the new variables $Z + N$ and $Z - N$, the substitution of Z and N by the new variables will not probably contribute to any significant improvement.
- Option 4: Inspired in the shell model and trying not to have redundancy in inputs. $FS(Z)$, $FS(N)$, $FO(Z)$, $FO(N)$, $Parity(Z)$ and

$Parity(N)$ normalized as inputs, and three feed-forward hidden layers of 11, 9 and 5 neurons, with $6 \cdot 11 + 11 \cdot 9 + 9 \cdot 5 + 5 \cdot 1 = 215$ connections.

- Option 5: Using all considered inputs, even with redundancy, except those of the semiempirical mass formula. Z, N, $Z+N$, $N-Z$, $FS(Z)$, $FS(N)$, Z_{fs}, N_{fs}, $FO(Z)$, $FO(N)$, $Parity(Z)$ and $Parity(N)$ normalized as inputs, and three feed-forward hidden layers of 11, 9 and 5 neurons, with $12 \cdot 11 + 11 \cdot 9 + 9 \cdot 5 + 5 \cdot 1 = 281$ connections. Therefore it makes use of the information of the liquid drop model, of the shell model and of the parity. If both the information of the liquid drop model and of the shell model can be used separately to build good models, using both together could allow a much better modeling, for each set of inputs could facilitate the modeling in a different manner and provide information that although being basically redundant, captures directly and without any further need of abstraction, different and relevant aspects of the problem.

- Option 6: Using all considered inputs, even the terms of the semiempirical mass formula. Z, N, $Z+N$, $N-Z$, $FS(Z)$, $FS(N)$, Z_{fs}, N_{fs}, $FO(Z)$, $FO(N)$, $Parity(Z)$, $Parity(N)$, T_2^{smf}, T_3^{smf}, T_4^{smf} and T_5^{smf} normalized as inputs, where T_i^{smf} stays for the i-th term of the semiempirical mass formula[11] given in equation (3), and three feed-forward hidden layers of 11, 9 and 5 neurons, with $16 \cdot 11 + 11 \cdot 9 + 9 \cdot 5 + 5 \cdot 1 = 325$ connections. The semiempirical mass formula is based on solid physical principles and offers a simple analytical model of the behavior of $B_{nucleus}$ with small error, which makes interesting combining it with an artificial neural network. Including each one of its terms offers more possibilities -and could allow a better modeling- than including their sum.

The number of connections considered does not seem excessive, since the number of training examples is much greater.

Regarding the neurons, a sigmoid activation function has been used for the hidden layers, an identity function for the input layer and an identity plus bias function for the output layer.

The learning algorithm used has been the usual backpropagation method with a small amount of momentum.

7 Results

Table 1 shows the mean square error and the root mean square error -the error that can be expected in predictions- for each trained ANN -measured

[11] The first term of the semiempirical mass formula has already been included with $Z + N$.

with the test set and applied the pertinent transform given in equation (7)- as well as for the semiempirical mass formula.

Option	Mean square error	Root mean square error (MeV)
smf	12.5577	3.54
1	7.0963	2.66
2	3.3195	1.82
3	2.2485	1.50
4	1.6297	1.28
5	0.39868	0.631
6	0.47653	0.690

Table 1: Errors of the trained ANNs and of the semiempirical mass formula (smf).

The simple ANN of the option 1 already improves the semiempirical mass formula. Figure 7 shows a projection graph of the output of the ANN as a function of the inputs N and Z (normalized). In figure 1 the prediction of $B_{nucleon}$ was plotted while here the prediction is of $B_{nucleus}$, but being the difference the multiplication by the number of nucleons $Z+N$, it is easy to realize that both agree.

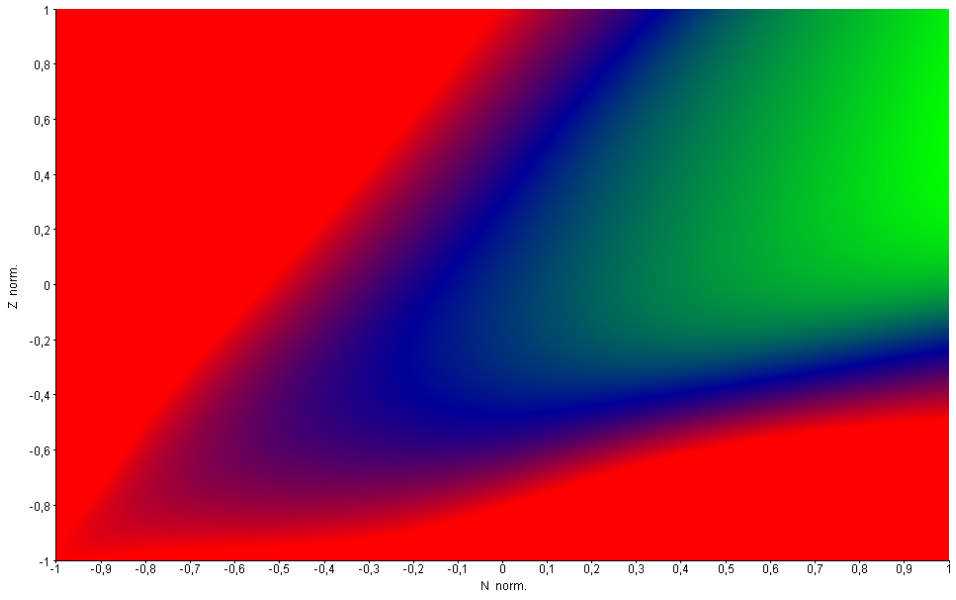

Figure 7: Projection graph of the output of the ANN of the option 1 as a function of its inputs.

The use of one additional hidden layer in option 2, with the consequent increase in the number of neurons and connections, substantially improves the results.

Option 3 shows a small improvement, probably due to the inclusion of the parity, although the increment in the number of connections could also be responsible in part.

The ANN of the option 4, which only has inputs corresponding to the shell model, slightly improves the results of the previous ANN -which has inputs of the liquid drop model and the parity.

The ANN of the option 5 is much better than the previous ones. Although the inputs have much redundancy, combining the inputs of the liquid drop model and the shell model results in a great improvement, which is not surprising, since both models are of importance when explaining nuclear properties.

Finally, the ANN of the option 6 -which has additional inputs corresponding to the terms of the semiempirical mass formula- does not improve the previous ANN. The small difference in the error can be due to the random initialization of the weights and the differences that arise in the training process.

It must be pointed out that the best ANN, that of the option 5, improves greatly the semiempirical mass formula -in a factor of 5 with a root mean square error of 0.631 MeV- achieving with 281 connections results as good as those achieved by good analytical models, while in [3] is stated that the lowest deviation to date (2003) obtained with a neural network was of 0.7 MeV. Our analysis so far shows that an appropriate decomposition of Z and N in the inputs is crucial for achieving these good results.

However, this ANN has two important disadvantages: being an ANN it acts like a black box that cannot be understood, and it has a lot of parameters, though many less than the number of training examples. It also has been observed that errors usually are greater in light atoms than in heavy atoms. This is not surprising for there are less light atoms than heavy atoms, and they are less regular and harder to model as can be seen in figures 3, 4 and 5. Typical errors are around 0.5 MeV in heavy atoms, but in some light atoms errors can be as high as 2 or 3 MeV.

Trying variations and refinements, the ANN of the option 5 could be still improved. It could be tested if any of the inputs can be removed without getting worse results, together with variations of the architecture with different number of neurons in hidden layers, of the activation functions and of the learning algorithm. It could also be tried to prune the network by removing less significant connections, leaving a network with less parameters without significantly increasing the error; removing in this way the irrelevant details of the network -while preserving the essential part of it- constitutes an abstraction process that could lead to a better model for making extrapolations.

8 Improved semiempirical mass formula

It would be interesting to get not an ANN, but an analytical model that allows its mathematical comprehension and sheds light to the underlying physical processes. Figure 6 shows that the error of the semiempirical mass formula can be modeled approximately by a second degree polynomial of the fraction of occuppation of the outer shell of protons $FO(Z)$ and of neutrons $FO(N)$. More exactly, the points in figure 6 seem to form a paraboloid centered in $(1/2, 1/2)$. Therefore, a simple improved model can be created by optimizing[12] the weighted sum of the terms of the semiempirical mass formula, the variables $(FO(Z)-1/2)^2$ and $(FO(N)-1/2)^2$, and a constant.

This optimization has been done with all 2228 examples, yielding the following improved semiempirical mass formula $B^{ismf}_{nucleus}$ with a mean square error of 8.0489 and a root mean square error of 2.84 MeV.

$$B^{ismf}_{nucleus} = a_v\,(Z+N) - a_s\,(Z+N)^{2/3} - a_c\,\frac{Z\,(Z-1)}{(Z+N)^{1/3}} - a_{sym}\,\frac{(N-Z)^2}{Z+N} + \delta(Z,N) + \\ + a_{foz}\left(FO(Z)-\frac{1}{2}\right)^2 + a_{fon}\left(FO(N)-\frac{1}{2}\right)^2 + b \tag{8}$$

where

$$\delta(Z,N) = \begin{cases} a_p\,(Z+N)^{-1/2} & \text{for } Z \text{ and } N \text{ even} \\ -a_p\,(Z+N)^{-1/2} & \text{for } Z \text{ and } N \text{ odd} \\ 0 & \text{for } Z+N \text{ odd} \end{cases} \tag{9}$$

and

$$FO(X) = \frac{X - X_{fs}}{OS(X)} \tag{10}$$

being X_{fs} the number of particles of the type indicated by X in filled shells, $OS(X)$ the number of particles of the type indicated by X that can contain the outer shell or its capacity, and $a_v = 15.596069$ MeV, $a_s = 17.735317$ MeV, $a_c = 0.70263955$ MeV, $a_{sym} = 22.754471$ MeV, $a_p = 10.005650$ MeV, $a_{foz} = 15.532$ MeV, $a_{fon} = 20.416$ MeV, $b = 0.554119$ MeV, the first five of which are, as expected, very similar to the constants of the semiempirical mass formula. It can also be appreciated that the values of a_{foz} and a_{fon} are similar, which can also be expected when observing figure 6 and recalling that the shell model is based on considerations about the nuclear force, which is equal for protons and neutrons.

[12] This optimization can be easily done using a very simple ANN without hidden layers.

Considering the potential that is supposed to describe the nuclear interaction, it can be deduced that magic numbers are given by the following formula,

$$Magic(n) = \begin{cases} \frac{n(n^2+3n+2)}{3} & \text{for } 3 \geq n \geq 1 \\ \\ \frac{n(n^2+5)}{3} & \text{for } n > 3 \end{cases} \tag{11}$$

due to which we can give expressions for X_{fs} and $OS(X)$ without any necessity of parameters,

$$X_{fs} = Magic(m) \;\; | \;\; Magic(m) \leq X < Magic(m+1) \tag{12}$$

$$OS(X) = Magic(m+1) - Magic(m) \;\; | \;\; Magic(m) \leq X < Magic(m+1) \tag{13}$$

and therefore[13]

$$FO(n) = \frac{n - Magic(m)}{Magic(m+1) - Magic(m)} \;\; | \;\; Magic(m) \leq n < Magic(m+1) \tag{14}$$

This new formula with 8 parameters is better than the semiempirical mass formula with 5, but the improvement is small, since the root mean square error only reduces from 3.54 MeV to 2.84 MeV. Despite of it, it is not negligible, because the improvement is about the double of the improvement that the inclusion of the pairing term -corresponding to the parity- produces. The improvement is modest due to differences between the concave surfaces of figure 3.

9 Conclusions

Adequately decomposing the inputs, ANNs can obtain models much better than the semiempirical mass formula, even comparable to good analytical models. Nevertheless, they have two disadvantages: because of having a lot of parameters they can be worse when extrapolating, and they act like black boxes that cannot be analytically understood.

[13]Here it must be taken into consideration the case of $Z \leq 1$ or $N \leq 1$, for wich we can define $Magic(0) = 0$ by extending the interval of the first case in equation (11) to $3 \geq n \geq 0$.

An obvious advantage of ANNs is that it is not necessary to have any understanding of the underlying physical processes, although it can help when decomposing the inputs. It suffices to use directly related inputs -in our case Z and N- to get a reasonably good model. Therefore they can be used in analogous cases where our knowledge is poor. For example, the mass of the hadrons could be modeled in base of their constituent quarks.

On the other hand, the good results obtained in this problem suggest that apart from the mass other nuclear properties that in principle are harder to model and predict could also be modeled with ANNs, e.g., radioactive decay modes and half-lives of nuclei.

ANNs not only are useful for making predictions, but they can also prove that there is an underlying order in the system, which can be very useful in the case of poorly explained physical systems in which they could discover and model relations.

Finally, we only have trained 6 ANNs, but training many more and trying variations and refinements, the error of 0.631 MeV would probably be reduced. Decomposing the inputs in ways other than the used here -with the help of more advanced nuclear models- could also provide further improvements.

References

[1] P. Roy Chowdhury and D. N. Basu, *Nuclear Matter Properties with the Re-evaluated Coefficients of Liquid Drop Model*, Acta Physica Polonica B, Vol. 37, No. 6 (2006), 1833-1846.

[2] K. S. Krane, *Introductory nuclear physics*, Wiley, 1987.

[3] D. Lunney, J.M. Pearson and C. Thibault, *Recent trends in the determination of nuclear masses*, Rev. Mod. Phys. 75 (2003), 1021-1082.

[4] A.H. Wapstra, G. Audi, and C. Thibault, *The AME2003 atomic mass evaluation (I). Evaluation of input data, adjustment procedures*, Nuclear Physics A, Vol. 729 (2003), 129-336.

[5] G. Audi, A.H. Wapstra, and C. Thibault, *The AME2003 atomic mass evaluation (II). Tables, graphs, and references*, Nuclear Physics A, Vol. 729 (2003), 337-676.